

PATENT ABSTRACTS OF JAPAN

(11) Publication number: **09146968 A**(43) Date of publication of application: **06.06.97**(51) Int. Cl. **G06F 17/30**(21) Application number: **07302463**(22) Date of filing: **21.11.95**(71) Applicant: **HITACHI LTD**(72) Inventor: **TADA KATSUMI
KATO KANJI
ASAKAWA SATOSHI
AZUMA AKIO**

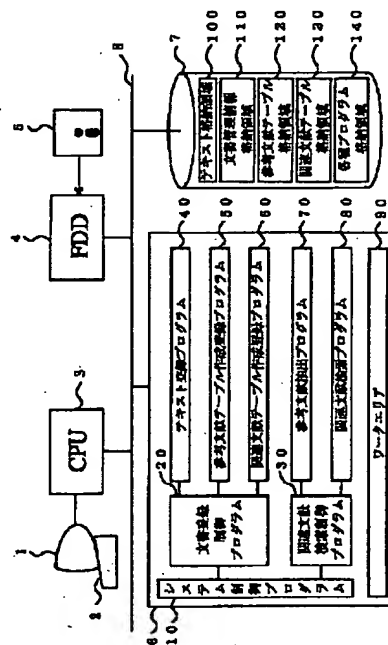
(54) DOCUMENT RETRIEVING METHOD

(57) Abstract:

PROBLEM TO BE SOLVED: To retrieve the latest relative document by regarding the same document that a user specifies as reference document and retrieving another document to be referred to.

SOLUTION: A test registration program 40 stores the text data of a registered document in a text storage area 100 on a magnetic disk 7, and also generates document management information for the registered document and stores it in the text storage area on the magnetic disk. Further, a reference document table generation and registration program 50 registers the document numbers of reference documents for respective registered documents to generate a reference document table. The generated reference document table is stored in a reference document table storage area 120 on the magnetic disk 7. Thus, the document obtained by referring to the relative document table generated in a relative document table generation and registration step for a reference document extracted in a reference document extraction step is extracted as a relative document.

COPYRIGHT: (C)1997,JPO



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平9-146968

(43) 公開日 平成9年(1997)6月6日

(51) Int. Cl. ⁶	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 F 17/30			G 0 6 F 15/419	3 2 0
			15/40	3 7 0 A
			15/401	3 2 0 C

審査請求 未請求 請求項の数1 O L (全 10 頁)

(21) 出願番号 特願平7-302463

(22) 出願日 平成7年(1995)11月21日

(71) 出願人 000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(72) 発明者 多田 勝己

神奈川県川崎市幸区鹿島田890番地の12

株式会社日立製作所情報・通信開発本部内

(72) 発明者 加藤 寛次

神奈川県川崎市幸区鹿島田890番地の12

株式会社日立製作所情報・通信開発本部内

(72) 発明者 浅川 悟志

神奈川県横浜市戸塚区戸塚町5030番地 株

式会社日立製作所ソフトウェア開発本部内

(74) 代理人 弁理士 小川 勝男

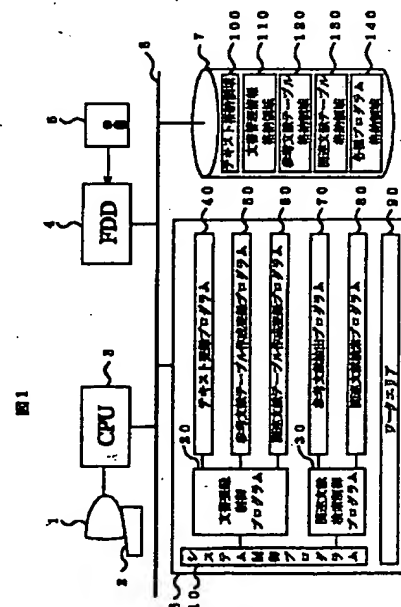
最終頁に続く

(54) 【発明の名称】 文書検索方法

(57) 【要約】 (修正有)

【課題】 ユーザが指定した文書と同じ文書を参考文献として参照している別の文書を検索することにより、ユーザの指定した文書に関連する最新の文献を検索する。

【解決手段】 文書情報をテキスト格納領域100に登録し、参考文献テーブル120及び関連文献テーブル130を作成し、作成されたテーブルを用いて関連のある参照文献を検索する。



【特許請求の範囲】

【請求項1】文書データベースに登録された複数の登録文書を検索する文書検索方法において、前記登録文書中で参照されている参考文献に対する当該登録文書からの第1のリンク情報を作成するとともに、前記登録文書に対する当該登録文書を参照している他の登録文書からの第2のリンク情報を作成しておき、使用者から登録文書が指定された場合、前記第1のリンク情報を用いて前記指定された登録文書の参考文献を前記文書データベースから抽出し、前記第2のリンク情報を用いて、前記抽出された参考文献に関連する文書を抽出することを特徴とする文書検索方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、文書データベースからユーザが指定した文書と関連性の高い文書を検索する文書検索方法に関する。データベース、文書管理システム、文書ファイリングシステムおよび電子図書館システムなどに適用されるものである。

【0002】

【従来の技術】論文などの技術文書を読み関心のある技術の調査を行う場合、その論文の中に引用されている参考文献を順次又引きしていく方法がよく用いられる。この方法では、図2に示すように引用元の文献1の中で参照されている参考文献1および参考文献2を検索し、その内容を読む。もし、その文書だけでは不十分な場合には、さらにその文書中で参照されている参考文献a、bを調べる。そして、所望の文書が得られるまで順次参考文献の又引きを繰り返す。

【0003】

【発明が解決しようとする課題】しかし、このような参考文献を又引きしていく方法では、最初に閲覧した文献より新しい文献を見つけることができない。すなわち、最新の関連文献を見つけることができないという問題がある。

【0004】図2に示す例で説明すると、文献1の中で参照されている参考文献1および参考文献2は文献1よりも古い文献である。また、に参考文献1の中で参照されている参考文献aおよび参考文献bは、参考文献1よりもさらに古い文献である。このように、又引きを繰り返す度に得られる文献は古い文献となるため、元の文書より新しい文献を得ることができず最新の文献を得ることができないという問題がある。

【0005】本発明が解決しようとする課題は、対象とする文献に関連する最新の文献を得ることのできる文書検索システムを提供することにある。

【0006】

【課題を解決するための手段】上記の課題を解決するために本発明は、以下の構成を採用する。

(1)登録文書をそのままテキストデータとして登録するテキスト登録ステップと、(2)該登録文書中で参照されている参考文献に対する該登録文書からのリンク情報を登録した参考文献テーブルを作成し登録する参考文献テーブル作成登録ステップと、(3)該登録文書に対する該登録文書を参照している文書からのリンク情報を登録した関連文献テーブルを作成し登録する関連文献テーブル作成登録ステップと、(4)関連文献の検索時には、該参考文献テーブル作成登録ステップで作成された参考文献テーブルを参照することにより、ユーザが指定した文書の中で参照されている参考文献を抽出する参考文献抽出ステップと、(5)該参考文献抽出ステップで抽出された参考文献に対し、該関連文献テーブル作成登録ステップで作成された関連文献テーブルを参照することにより得られた文書を、関連文献として抽出する関連文献抽出ステップを有する文書検索方法である。

【0007】以上の構成である文書検索方法を用いることにより、ユーザが指定した文書と同じ文書を参考文献として参照している別の文書を検索することができる。一般に、参考文献としては各分野を代表する文献が挙げられることが多く、最新の文書についても古い文書と同じ文書を参考文献として挙げる可能性が高いため、参考文献をキーとしてユーザが指定した文書と関連性の高い最新の文書を検索することができることになる。

【0008】以下、図3を用いて本発明の概要について説明する。

【0009】文書の登録時には、まずテキスト登録ステップで文書1、文書2、文書3および文書4をテキストデータとして登録する。

【0010】次に、参考文献テーブル作成登録ステップでは、各登録文書中で参照されている参考文献の番号を登録した参考文献テーブルを作成し登録する。すなわち、図3に示す例では、文書1で参考文献として文書2を参照しているため文書1から文書2に参照先リンクを生成する。文書2についても同様に文書5を参照しているため、文書2から文書5に参照先リンクを生成する。以下、文書3および文書4についても同様の処理を繰り返すことにより図4に示す参照先リンク情報を生成する。そして、参照先リンク情報をもとに、各登録文書中で参照されている参考文献の番号を登録した参考文献テーブルを作成する。

【0011】さらに、関連文献テーブル作成登録ステップでは、各登録文書についてその文書を参照している参照元文書からのリンク情報を登録した関連文献テーブルを作成し登録する。例えば、図4において参照先リンク情報を逆向きに参照する。すなわち、参照先リンク情報の参照元エントリと参照先エントリを入れ替え、その結果を参照先文書番号でソーティングすることにより、各文書を参照している参照元文書に対する参照元リンク情報を生成する。そして、参照元リンク情報をもとに各登

録文書について、その文書を参照している参照元文書の番号を登録した関連文献テーブルを作成する。

【0012】以上が登録時の処理である。

【0013】次に検索時には、以下の処理を施す。

【0014】参考文献抽出ステップで指定された文書に対し参考文献テーブルを参照することにより、ユーザが指定した文書の中で参照している参考文献の文書番号を取得する。すなわち、図5に示す例ではユーザが指定した文書1に対して参考文献テーブルを参照することにより文書2が参考文献として得られる。

【0015】そして、関連文献抽出ステップでは参考文献抽出ステップで得られた参考文献に対し関連文献テーブルを参照することにより、ユーザが指定したと関連性の高い文書を抽出する。すなわち、図5に示す例では参考文献抽出ステップで得られた文書2に対し関連文献テーブルを参照することによって、文献2を参照している文書として文書1、文書3および文書4を得る。

【0016】以上の処理によって、文書1と関連性の高い文書として文書3と文書4を得ることができる。

【0017】このように本発明では、従来の方法では得られなかった元の文書よりも新しい文書が得ることができる。このため、指定された文書と関連性の高い最新の文書の検索が可能になる。

【0018】

【発明の実施の形態】本発明の文書検索方法を実現する1構成である検索システムを図1に示す。本図に示す検索システムは、検索結果を表示するディスプレイ1、登録および検索のコマンドを入力するキーボード2、登録処理および検索処理を実行する中央演算処理装置CPU3、フロッピディスクからデータを読み出すフロッピディスクドライバFDD4、データベースへ登録する文書データを格納したフロッピディスク5、登録および検索用のプログラムならびにデータなどを一時的に格納する主メモリ6、検索用データを格納する磁気ディスク7およびこれらを接続するバス8で構成される。

【0019】主メモリ6には、システム制御プログラム10、文書登録制御プログラム20、関連文献検索制御プログラム30、テキスト登録プログラム40、参考文献テーブル作成登録プログラム50、関連文献テーブル作成登録プログラム60、参考文献抽出プログラム70および関連文献検索プログラム80が磁気ディスク7から読み出されるとともにワークエリア90が確保される。

【0020】また、磁気ディスク7にはテキスト格納領域100、文書管理情報格納領域110、参考文献テーブル格納領域120、関連文献テーブル格納領域130、各種プログラム格納領域140が確保されている。なお、本実施例ではこれらの格納領域を磁気ディスク7上に確保したが、光磁気ディスク装置など他の二次記憶装置であっても構わない。以上が本検索システムの構成

である。

【0021】次に、本実施例に示す文書検索方法の文書登録時の処理と検索時の処理について説明する。

【0022】始めに、キーボード2から入力される登録コマンドによりシステム制御プログラム10は、文書登録制御プログラム20を起動して、文書の登録処理を開始する。この文書登録時の処理の概略を図6に示すPAD図を用いて説明する。

【0023】始めに、文書登録制御プログラム20はステップ1000でテキスト登録プログラム40を起動する。テキスト登録プログラム40では、登録文書のテキストデータを磁気ディスク7内のテキスト格納領域100に格納するとともに、登録文書に対し文書管理情報を作成しこれを磁気ディスク内のテキスト格納領域110に格納する。

【0024】次に、テキスト登録制御プログラム20は、ステップ1001で参考文献テーブル作成登録プログラム50を起動する。参考文献テーブル作成登録プログラム50では、各登録文書の参考文献の文書番号を登録することにより参考文献テーブルを作成する。作成された参考文献テーブルを磁気ディスク7上の参考文献テーブル格納領域120に格納する。

【0025】最後に、テキスト登録制御プログラム20は、ステップ1002で関連文献テーブル作成登録プログラム60を起動する。関連文献テーブル作成登録プログラム60では、各登録文書について、以下の処理を行う。登録文書を参照している参照元文書の番号を登録することにより関連文献テーブルを作成して、磁気ディスク7上の関連文献テーブル格納領域に格納する。以上が、登録時処理の概要である。

【0026】次に、上記登録プログラムの処理内容について、図3に示す文書が登録された場合を例に説明する。

【0027】始めに、テキスト登録プログラム40で、フロッピディスクフロッピディスクドライバ4に挿入されたフロッピディスク5から文書1、文書2、文書3、文書4および文書5のテキストデータをワークエリア90に読み込む。このことは、図7のステップ1100に示す通りである。そして、ステップ1101で、ワークエリア90に読み込まれたテキストデータを磁気ディスク7内のテキスト格納領域100へ格納する。

【0028】さらに、ステップ1102で、磁気ディスク7内の文書管理情報格納領域110から文書管理情報を主メモリ6上のワークエリア90に読み込む。なお、本実施例では、登録文書を管理する情報として図8に示す対応表を用いる。

【0029】そして、ステップ1103で、キーボードから登録文書のタイトルを入力させる。

【0030】そして、ステップ1104で、入力されたタイトルを以下のものと併せて格納する。併せて格納す

るものは、登録文書の文書番号および該当文書を格納するファイル名である。このことにより、文書管理情報を追加登録する。

【0031】そして、以上の処理が終了するとステップ1105で、文書管理情報を磁気ディスク7内の文書管理情報格納領域110に格納し、テキスト登録プログラム20の処理を終了する。

【0032】なお、登録文書はフロッピディスクを用いて入力するだけでなく通信回線（図1に示していない）等を用いて他の装置から入力されるような構成をとって10 もかまわない。この場合、通信制御装置等を用いてバス8を通して入力させる。

【0033】次に、参考文献テーブル作成登録プログラム50で、まずステップ1200において各登録文書内で参照している参考文献のタイトルをキーボード2から入力する。このことは、図9に示す。例えば文書番号1の文書については、参考文献の文書名として「通信会議システム××の開発」が入力される。

【0034】そしてステップ1201で、参考文献のタイトルで文書管理情報内のタイトルを検索する。このことにより各参考文献の文書番号を取得する。すなわち、20 文書番号1の文書については、参考文献のタイトルである「通信会議システム××の開発」で文書管理情報内のタイトルを検索することにより参考文献の文書番号として文書番号2を得る。

【0035】さらに、ステップ1202で、各登録文書と参考文献の文書番号をリンク情報を登録することにより、参照先リンクを生成する。すなわち、文書番号1の文書からは参照先リンクとして文書番号2が生成される。このため、図3に示す文書に対して図4に示す参照30 元リンク情報が、生成されることになる。

【0036】そして、ステップ1203で、ステップ1202において生成された参照先リンク情報をもとに、各登録文書中で参照されている参考文献の番号を登録する。このことにより図4に示す参考文献テーブルを作成する。

【0037】そして、以上の処理が終了すると以下の処理を行う。ステップ1204で、参考文献テーブルと参照先リンク情報を磁気ディスク7内の参考文献テーブル格納領域120に格納する。こうして、参考文献テーブル作成登録プログラム50を終了する。40

【0038】最後に、関連文献テーブル作成登録プログラムでは、図10に示す通り以下の処理を行う。まずステップ1300で、参照先リンク情報を磁気ディスク7内の参考文献テーブル格納領域120から主メモリ6上のワークエリア90に読み込む。

【0039】そしてステップ1301で、参照先リンク情報の参照先のエントリと参照元のエントリを入れ替える。さらに、参照元文書番号でソーティングすることにより参照元リンク情報を生成する。このことは、図4に50

示す通りである。また、参照元リンク情報をもとに、各登録文書中で参照されている参考文献の番号を登録する。このことにより参考文献テーブルをステップ1302で作成する。

【0040】そして、以上の処理が終了するとステップ1303で、関連文献テーブルと参照元リンク情報を磁気ディスク7内の関連文献テーブル格納領域130に格納する。こうして、関連文献テーブル作成登録プログラム60を終了する。以上が、登録時処理の具体的な処理内容である。

【0041】次に、検索時の処理手順を図11に示す。以下、図5に示す文書に対して文書番号1に関連した文書を検索する場合を例に検索時の処理を説明する。

【0042】始めに、キーボード2から入力される検索コマンドによりシステム制御プログラム10は、関連文献検索制御プログラム30を起動して、関連文献検索処理を開始する。

【0043】そして、関連文献検索制御プログラム30は、ステップ2000で参考文献抽出プログラム70を起動する。参考文献抽出プログラム70では、参考文献テーブルを参照することによりユーザが指定した文書の中で参照している参考文献の文書番号を取得する。すなわち、図5の例ではユーザが指定した文書番号1に対し、参考文献テーブルを参照する。このことにより、文書番号1の中で参照している参考文献の番号として文書番号2を得る。

【0044】次に、関連文献検索制御プログラム30は、ステップ2001で関連文献抽出プログラム80を起動する。関連文献抽出プログラム80では、参考文献抽出プログラム70で抽出された参考文献の文書番号をもとに関連文献テーブルを参照する。このことにより、ユーザが指定した文書中の参考文献を参照している文書の番号を取得する。すなわち、図5に示す例では、ユーザの指定した文書番号1の文書に対する参考文献として得られた文書番号2で関連文献テーブルを参照する。このことにより、文書番号1、3および4を得る。以上の処理によって、文書1と関連性の高い文書として文書3と文書4を得ることができる。以上が、検索時処理の具体的な内容である。

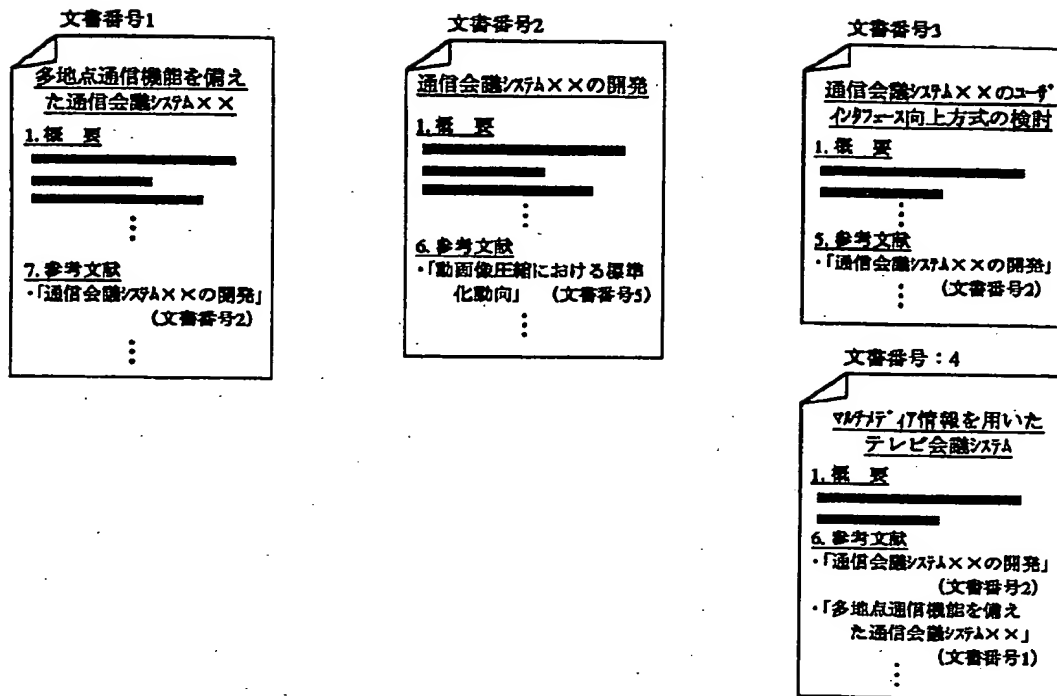
【0045】以上示したように、本発明によればユーザが指定した文書と同じ文書を参考文献として参照している別の文書を検索する。したがって、従来の関連文書検索方法では検索することのできなかった最新の関連文献の検索を実現することができる。

【0046】

【発明の効果】本発明によれば、ユーザが指定した文書と同じ文書を参考文献として参照している別の文書を検索することにより、従来の関連文書検索方法では検索することのできなかった最新の関連文献の検索を実現することが可能となる。

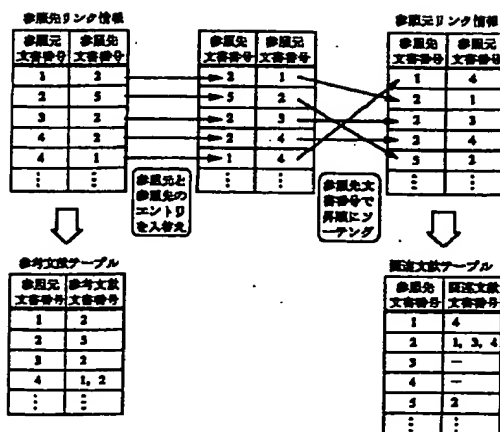
【図3】

図 3

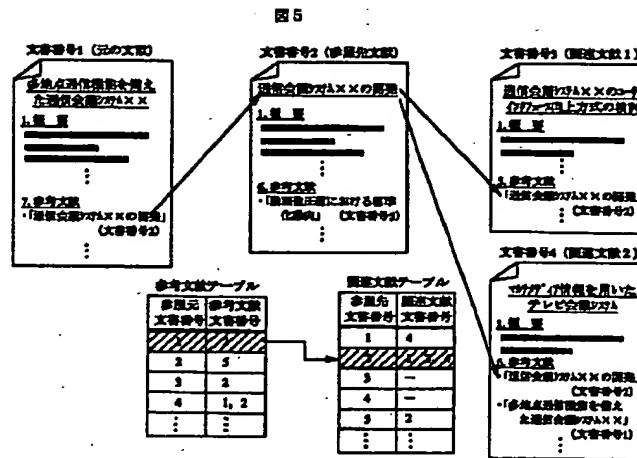


【図4】

図 4

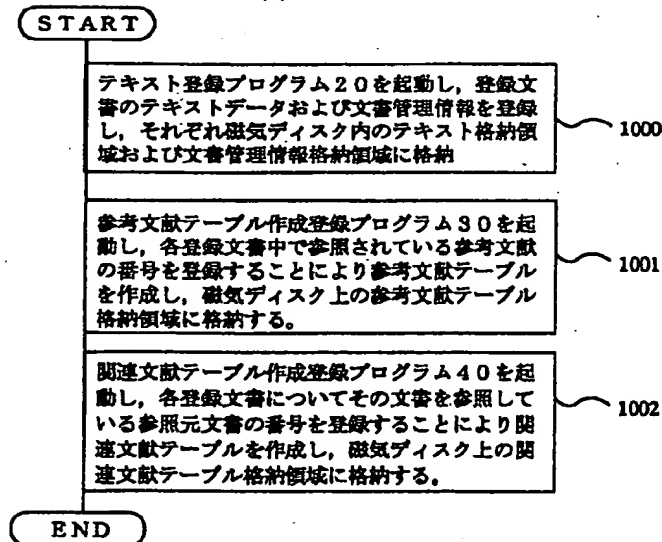


【図5】



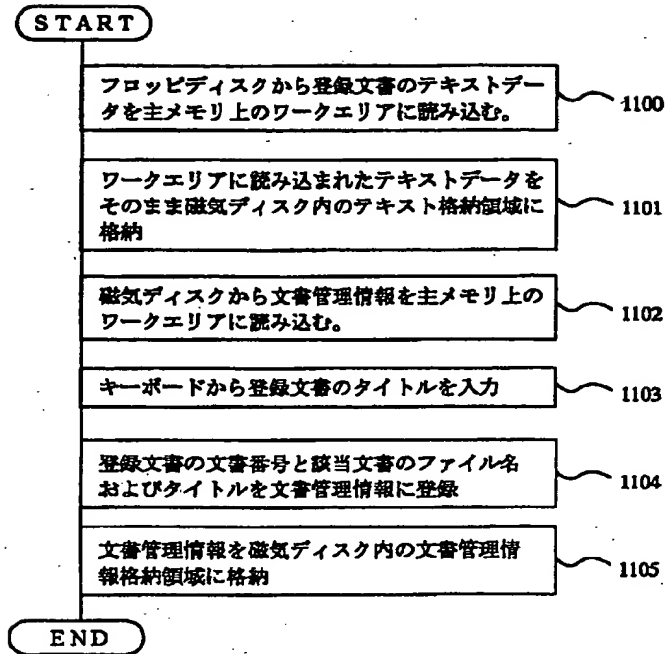
【図6】

図6



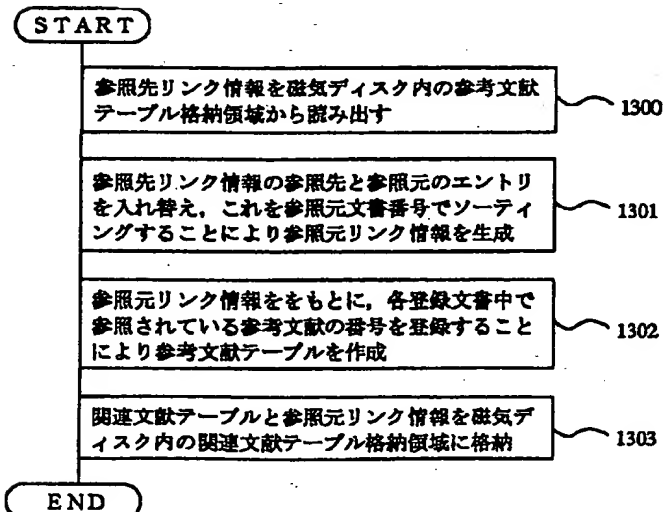
【図7】

図7

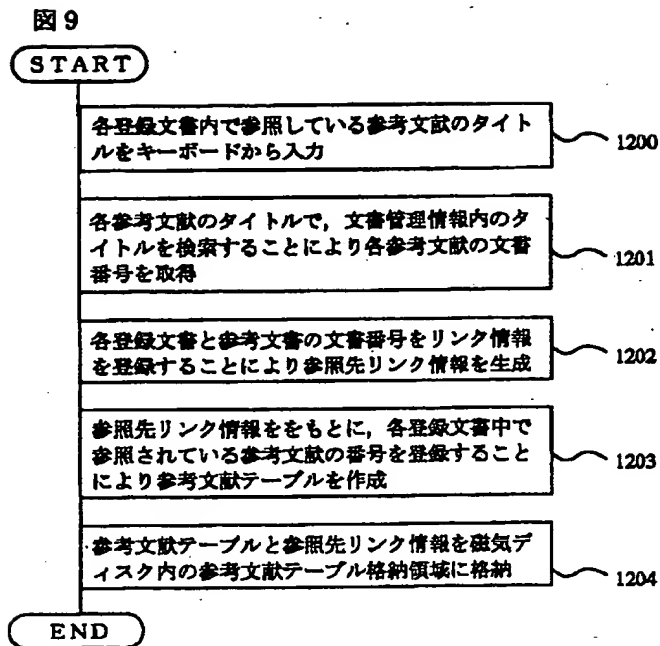


【図10】

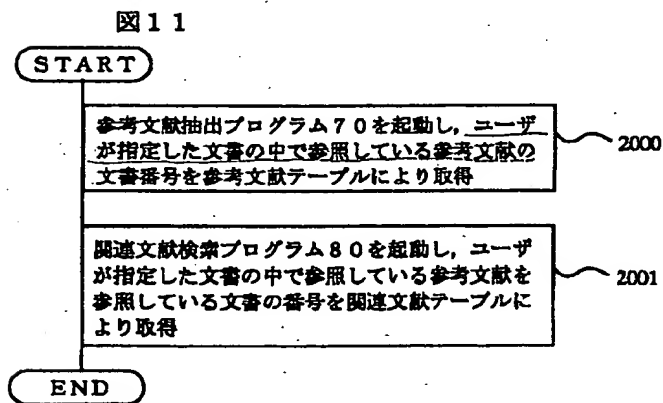
図10



【図9】



【図11】



フロントページの続き

(72)発明者 東 秋夫

東京都江東区新砂一丁目6番27号 株式会
社日立製作所公共情報事業部内